GATE LYON SAINT-ÉTIENNE
UMR 5824

93, chemin des Mouilles
69130 Ecully - France

Maison de l'Université, Bâtiment B
10, rue Tréfilerie
42023 Saint-Etienne cedex 02 - France

http://www.gate.cnrs.fr
gate@gate.cnrs.fr

WP 2037 – December 2020

Productive Workfare? Evidence from Ethiopia's Productive Safety Net Program

Jules Gazeaud, Victor Stephane

Abstract:

Despite the popularity of public works programs in developing countries, there is virtually no evidence on the value of the infrastructure they generate. This paper attempts to start filling this gap in the context of the PSNP – a largescale program implemented in Ethiopia since 2005. Under the program, millions of beneficiaries received social transfers conditional on their participation in activities such as land improvements and soil and water conservation measures. We examine the value of these activities using a satellite-based indicator of agricultural productivity and (reweighted) difference-in-differences estimates. Results show that the program is associated with limited changes in agricultural productivity. The upper bound of the main estimate is equivalent to a 3.6 percent increase in agricultural productivity. This contrasts with existing narratives and calls for more research on the productive effects of public works.

Keywords:

Social Protection, Public Works, Transfers, Ethiopia, PSNP

JEL codes:
I38, O13, O22, Q15

Productive Workfare? Evidence from Ethiopia's Productive Safety Net Program*

Jules Gazeaud † Victor Stephane ‡

December 2020

Abstract

Despite the popularity of public works programs in developing countries, there is virtually no evidence on the value of the infrastructure they generate. This paper attempts to start filling this gap in the context of the PSNP – a large-scale program implemented in Ethiopia since 2005. Under the program, millions of beneficiaries received social transfers conditional on their participation in activities such as land improvements and soil and water conservation measures. We examine the value of these activities using a satellite-based indicator of agricultural productivity and (reweighted) difference-in-differences estimates. Results show that the program is associated with limited changes in agricultural productivity. The upper bound of the main estimate is equivalent to a 3.6 percent increase in agricultural productivity. This contrasts with existing narratives and calls for more research on the productive effects of public works.

Keywords: Social Protection; Public Works; Transfers; Ethiopia; PSNP

JEL Codes: I38; O13; O22; Q15

*This paper has benefited from the thoughts of Chris Barrett, Hugues Champeaux, Doug Gollin, Flore Gubert, Pascal Jaupart, Elisabeth Sadoulet, Olivier Santoni, Giulio Schinaia, as well as audiences at the 2020 NEUDC Conference and the International Workshop on Spatial Econometrics and Statistics. Gazeaud acknowledges the support received from Fundação para a Ciência e a Tecnologia (UID/ECO/00124/2013 and Social Sciences DataLab, Project 22209), POR Lisboa (LISBOA-01-0145-FEDER-007722) and POR Norte. The usual disclaimers apply.

†Gazeaud: NOVAFRICA, Nova School of Business and Economics, Universidade Nova de Lisboa, Campus de Carcavelos, Rua da Holanda 1, 2775-405 Carcavelos, Portugal, email: jules.gazeaud@novasbe.pt.

‡Stephane: Univ Lyon, UJM Saint-Etienne, GATE UMR 5824, F-42023 Saint-Etienne, France, email: victorstephane2@gmail.com.

"Over the years, the Productive Safety Net Programme has reduced soil loss by between 40-53%, increased land productivity by up to 400%, contributed to improved quality and flow of water, and decreased the damage from seasonal flooding." H.E. Ato Sileshi Getahun, Former Ethiopian State Minister of Agriculture (in Foreword of World Bank, 2013).

1 Introduction

Public Works Programs (PWP – also often referred to as workfare or cash-for-work programs) have become increasingly popular in developing countries, being implemented in Argentina, Ethiopia, India, and South Africa, among many others, and covering tens of millions of participants (Gehrke and Hartwig, 2018). They consist in providing short-term employment opportunities to poor and underemployed individuals in labor-intensive infrastructure projects. In doing so, they follow the dual objective to reduce poverty by transferring income to the poor and to use the workforce of participants to build infrastructure and public goods in the targeted areas. While they were traditionally used as crisis relief, they have recently triggered attention for broader purposes, including the need to cope with environmental degradation and climate change vulnerability (Subbarao et al., 2012).

A drawback of PWP lies in their implementation costs which are generally higher than for other poverty alleviation programs. Gehrke and Hartwig (2018) estimate that for each dollar spent, an average of 31 cents reaches beneficiaries in PWP against 42 cents in basic cash transfer programs. In addition, labor requirements in PWP usually imply welfare losses for beneficiaries including forgone income and unpleasant work (Murgai et al., 2015; Alik-Lagrange and Ravallion, 2018). These downsides are often justified by the assumption that PWP infrastructure generates important productive effects.[1] For instance, Subbarao et al. (2012) argue that "*there is no reason to do public works if the public goods generated do not have a positive impact on the community*". Yet, evidence on such aspects is particularly scant, preventing comprehensive cost-effectiveness exercises and comparisons with other poverty alleviation programs.

The present paper aims to start filling this gap by investigating the productive effects of

[1]Other arguments have been made to justify PWP, including self-targeting gains in the presence of asymmetric information (Besley and Coate, 1992; Dutta et al., 2014; Alik-Lagrange and Ravallion, 2015), improvements in income generating capacities through on-the-job or formal training (Gehrke and Hartwig, 2018), or insurance provision against risk (Alderman and Yemtsov, 2014; Gehrke, 2019). While we recognize the importance of these aspects, in this article we focus on the infrastructure generated by PWP because it is arguably a crucial – and yet unexplored – input to justify cost-effectiveness. In addition, according to Murgai et al. (2015), "*for the same budget, unproductive workfare has less impact on poverty than either a basic-income scheme or transfers tied to the government's assignment of ration cards. The productivity of workfare is thus crucial to its justification as an antipoverty policy*". See Ravallion (2019) for further discussion of the arguments for and against PWP.

the infrastructure generated in the context of the Productive Safety Net Program (PSNP). The PSNP is a flagship PWP implemented in Ethiopia since 2005. It provides social transfers to millions of beneficiaries in chronically food insecure woredas in exchange for their participation in labor-intensive activities. As most activities focus on land management and environmental projects such as soil and water conservation activities, the PSNP is sometimes considered as Africa's largest climate change adaptation program (Subbarao et al., 2012). An influential view, as illustrated in the introductory quote, is that the PSNP has reduced soil erosion and (drastically) improved land productivity. However, evidence on these aspects is sorely lacking.[2]

Assessing the productive effects of PSNP works is particularly challenging due to the lack of survey data over a sufficient period of time. We tackle this issue by relying on high resolution satellite data. More precisely, we build an indicator of agricultural productivity by combining the Normalized Difference Vegetation Index (NDVI) with highly disaggregated information on land use, crop types, and crop calendars. We show that this indicator is a good predictor of agricultural output. In addition, we enrich our dataset with data on climatic conditions (rainfall and temperature), topographic characteristics, night-time lights, and population density to control for a maximum of potential confounding factors. This results into a unique and comprehensive dataset providing information at the woreda-year level for the whole country over the 2000-2013 period.

We estimate the impact of the PSNP on agricultural productivity using a difference-in-differences estimator. As the PSNP was targeted rather than randomly allocated, finding a credible counterfactual is difficult. To overcome this challenge, we employ the inverse probability weighting method (Hirano et al., 2003). We show that this empirical strategy not only allows us to absorb differences in pre-treatment trends between treated and control woredas, but also to achieve more balanced groups. In addition, as ecological disparities across woredas may potentially lead to heterogeneous effects, we run separate regressions for highlands and lowlands. Last but not least, because the PSNP was not homogeneously implemented, some woredas being more treated than others, we use two additional treatment variables (in addition to a basic binary variable), namely the percentage of households treated in each woreda (*treatment intensity*), and the number of treated households per square kilometers (*treatment density*).

Our results provide no evidence that public works had a sizeable impact on agricultural

[2]We have been unable to find sources for the figures outlined in the introductory quote. These figures are mentioned in the Foreword of World Bank (2013) and also on page 27 of the same report: "*PSNP public works have improved the potential for growing food, which builds community resilience. Evidence shows a 40-53% reduction in sediment reaching streams through soil erosion, a 3-4 fold increase in land productivity, and many reports of higher crop production*". However, the report does not include references.

productivity. This conclusion holds regardless of the treatment variable used, suggesting that neither the least treated nor the least densely populated woredas drive these null estimates. Importantly, we provide evidence that these null effects are not explained by a lack of statistical power of our estimations. The upper bound of the main estimate is equivalent to a 3.6 percent increase in agricultural productivity and we estimate a minimum detectable effect size of at most 0.07 SD. Because benefits of public works could take time to manifest, we investigate the impact of the PSNP on a yearly basis and confirm the lack of visible effect even seven years after the beginning of the program. Finally, positive effects might only be visible in case of unfavorable climatic conditions. We estimate the impact of the program conditional on the level of rainfall and find no significant interactions.

To assess the validity of these results, we conduct several robustness checks. First, to check that omitted variables do not confound our results, we add economic and demographic control variables, and we compute Oster bounds (Oster, 2019). In both cases, coefficients hardly change and remain non-significant. Second, we restrict our sample to the common support as well as to the most treated woredas. Results are consistent with our baseline specification and show no significant effects. We also conduct estimates by elevation deciles instead of using an arbitrary cut-off to distinguish highlands and lowlands. Finally, we provide suggestive evidence that the null effects are not driven by selective migration patterns or by negative effects of the transfers on agricultural activities.

This paper contributes to the literature on the effects of PWP in general, and on the effects of the PSNP more precisely. This literature suggests that PSNP transfers have positive effects on various outcomes such as food security (Gilligan et al., 2009; Nega et al., 2010; Berhane et al., 2014), children nutritional status and human capital accumulation (Debela et al., 2015; Porter and Goyal, 2016; Favara et al., 2019; Mendola and Negasi, 2019), investments in technology adoption or tree planting (Andersson et al., 2011; Adimassu and Kessler, 2015; Alem and Broussard, 2018; Araya, 2020), while they do not seem to divert children from schooling or to increase child labor (Hoddinott et al., 2010) – two typical concerns with PWP.[3,4] However, to the best of our knowledge, the productive impacts of the PSNP remain largely unexplored. A noticeable exception is Filipski et al. (2016). Using a panel GMM estimator, the authors find that soil and water conservation measures increased the average yields of grain crops by about

[3]In other contexts, the available evidence on PWP effects is more mixed. See for example Ravi and Engler (2015), Rosas and Sabarwal (2016), Bertrand et al. (2017), Beegle et al. (2017), Maity (2020) on food security, Shah and Steinberg (2019), Ajefu and Abiona (2019), Li and Sekhri (2020) on human capital accumulation, and Galasso and Ravallion (2004), Ravallion et al. (2005), Zimmermann (2012), Imbert and Papp (2015), Deininger et al. (2016), Berg et al. (2018), Merfeld (2019) on labor market outcomes.

[4]These positive effects are regularly observed, generally at lower costs, in alternative poverty alleviation programs such as unconditional cash transfers (Baird et al., 2014; Haushofer and Shapiro, 2016; Ravallion, 2019).

2.8 percentage points but had no effect on non-grain crops. These results are, however, subject to the typical concerns about the use of GMM estimators to achieve causal inference (Roodman, 2009), and, as mentioned by the authors, some results could reflect a lack of statistical power.[5] Moreover, even in different contexts, the only study on the value of PWP infrastructure we are aware of is Christian et al. (2015). Using a randomized controlled trial of the Labor Intensive Works Program (LIWP) in Yemen, the authors show that water-related projects had large and positive effects on water accessibility in villages with poor baseline access. However, a clear concern in this study is that the results were derived from the subset of villages with completed projects at the time of the follow-up survey and could therefore reflect convergence in water access rather than the effects of the LIWP.[6] By combining an original and comprehensive dataset with a credible identification strategy, we address the aforementioned shortcomings and add to the effort to estimate the causal effect of PWP.

In addition, we provide another example of geospatial impact evaluations (BenYishay et al., 2017; Lech et al., 2018), and further substantiate their potential to study important questions at low costs. In particular, we add to recent studies relying on satellite data to measure agricultural yields (Burke and Lobell, 2017; Lobell et al., 2020). Combining NDVI data and georeferenced data on land use, crop types and crop calendars, we show the scope for constructing an indicator that is highly predictive of variations in survey-based agricultural output. This indicator offers great potential for assessing variation in yields at a large scale and high frequency. In the context of this study, it allowed us to recover multiple rounds of data prior to the implementation of the PSNP,[7] which has been particularly useful to test the parallel trends assumption prior to the program (a key assumption for our identification strategy).

The remainder of the paper is organized as follows. Section 2 provides background information on the program. Section 3 presents the data. Section 4 outlines the empirical strategy. Section 5 presents the results, and Section 6 concludes.

2 Background

With more than 100 million inhabitants, Ethiopia has currently the second largest population in Africa after Nigeria. The country is administratively divided into ethnically based regions, which are themselves subdivided into zones, and woredas (districts). Geographically, it is

[5]Lack of statistical power increases not only the likelihood of false negative but also that of false positive (Gelman and Carlin, 2014).

[6]Only 8 out of the 82 projects were completed at follow-up.

[7]To the best of our knowledge, other studies on the impact of the PSNP include at most one survey round prior to the implementation of the program.

composed of a vast territory made of mountains and plateaus lying at elevations above 1500m, divided by the Rift Valley, and surrounded by lowlands. Livelihood predominately depends on crop production in highlands and on agro-pastoralism in lowlands. Over the last decades, Ethiopia has faced severe droughts which often resulted in large-scale food crisis.[8] Despite some progress, it is still one of the poorest country in the world with a per capita income of US$772 in 2018 (World Bank data). Poverty is especially widespread in rural areas where most people are engaged in subsistence agriculture and face important environmental degradation. Environmental degradation not only reduces land productivity, but also the ability to manage water resources effectively. According to a World Bank report, "*[Ethiopian] land base has been damaged through erosion and degradation, land productivity has declined, and rainfall infiltration has reduced such that many spring and stream sources have disappeared or are no longer perennial*" (World Bank, 2006, p.2). Environmental degradation is particularly prevailing in Ethiopian highlands because of the steep slopes and widespread deforestation.[9]

The Productive Safety Net Program (PSNP) was launched in 2005 by the Government of Ethiopia in an attempt to provide a long-term solution to rural chronic food insecurity. It is among the largest public works programs to date. The PSNP replaced an old system where food aid depended on emergency humanitarian appeals for international assistance. Food aid was mainly delivered as emergency food-for-work programs (Porter and Goyal, 2016), and, according to the government's policy, no able-bodied person should receive food aid without working on a community project in return (Jayne et al., 2002; Quisumbing and Yohannes, 2005).[10] This system proved inefficient as assistance was unpredictable both for planners and local populations (Kehler, 2004; Rahmato, 2013). In contrast, the PSNP aimed to provide a predictable and reliable safety net to address chronic food insecurity and mitigate recurrent climate shocks. The PSNP was quickly rolled-out to reach approximately 8 million beneficiaries in 2006/07,[11] thereby becoming the largest workfare program in Africa and one of the largest

[8]The most prominent example is probably the 1983-1984 famine from which up to one million people are estimated to have died (Devereux, 2000).

[9]For a rather old but enlightening examination of environmental degradation in Ethiopian highlands, see a study commissioned by the FAO which argues that "*the highlands of Ethiopia contain what is probably one of the largest areas of ecological degradation in Africa, if not in the world*" (Hurni, 1983, p.ii).

[10]Prior public works could threaten the internal validity of our study if they were targeted towards the same areas and activities as the PSNP. While our data do not allow us to fully rule out this concern, we believe that it remains limited because aid receipts prior to the PSNP display a large degree of heterogeneity across woredas and do not overlap with assignment to the PSNP (Schinaia, 2016). In addition, there is evidence that the structures created prior to the PSNP were of poor quality and often rejected or removed by farmers and local communities because of the top-down implementation of activities (Shiferaw and Holden, 1998; Holden et al., 2006). If anything, the PSNP therefore represents a major breakthrough with respect to prior public works programs.

[11]Throughout the paper, we follow Gilligan et al. (2009) and Porter and Goyal (2016) who consider 2006 as a pretreatment year because the program was not fully rolled-out. Results are similar considering 2006 as a treatment year.

safety net programs. Today, it operates with an annual budget of more than US$500 million.

The main component of the PSNP consists of cash or food transfers to selected poor households conditional on their participation in labor-intensive projects. The program targets chronically food insecure households in chronically food insecure woredas using a mix of geographic and community-based targeting devices. The government identified chronically food insecure woredas based on the number of years they had required food assistance prior to 2005. Then, in each eligible woreda, local community councils known as food security task forces (FSTF) identified food insecure households, that is households that (i) have repeatedly faced food gaps or received food aid in the past three years, (ii) have suffered from a severe loss of assets due to a severe shock, and (iii) with no other source of support (family or social protection programs). These targeting guidelines were intended as a broad national framework but in practice the program allowed for regional and local adaptation by FSTF (Sharp et al., 2006). Able-bodied adults of beneficiary households could participate in public works only during the agricultural off-season in order to limit interference with farming and other income-generating activities.[12] The wage rate was initially set at 6 birr per day – approximately US$0.70 using the 2005 official exchange rate – and gradually increased in an effort to reflect inflation patterns (Sabates-Wheeler and Devereux, 2010; Hirvonen and Hoddinott, 2020). According to administrative data, PSNP activities generated 227 millions person-days of employment in 2008 (World Bank, 2016).

Most public works in the PSNP focus on watershed development, with the objectives to achieve environmental rehabilitation and increase agricultural productivity and resilience to climate shocks. Activities include the construction of area closure, terraces and flood control structures, agroforestry, gully control, or the renovation of traditional water bodies activities (see Table A1 for a detailed list of the activities undertaken between 2007 and 2009). Projects were selected locally through a community-based participatory approach and integrated into woredas development plans. The peculiar conditions found in pastoral regions (Afar and Somali) caused implementation delays and required some tweaking in terms of program design. In particular, these regions only started to benefit from public work activities in 2010 – beneficiary households received unconditional transfers until 2009/10. Our empirical strategy takes this specificity into account (see Section 4).

[12]Beneficiary households with no able-bodied adult members were included in the direct support component of the PSNP (i.e. unconditional transfers of the same amount). Direct support beneficiaries represent about 16 percent of total beneficiaries.

3 Data

To estimate the impact of PSNP infrastructure on crop production, we assemble an original database covering Ethiopia over the 2000-2013 period.[13] We rely on geo-referenced administrative data and high resolution satellite data to conduct our study at the woreda-year level.[14] This section describes the details of how we build this dataset.

3.1 Crop Production

We build an indicator of agricultural output by combining the Normalized Difference Vegetation Index (NDVI) with highly disaggregated information on land use, crop types, and crop calendars. NDVI data are available bi-monthly at a resolution of 250m and are widely used to proxy vegetation health and crop productivity (see e.g. Pettorelli et al., 2005; Wang et al., 2005; Atzberger, 2013; Klisch and Atzberger, 2016; Jensen et al., 2019; Ali et al., 2020; Asher and Novosad, 2020). Since our analysis are conducted at the woreda-year level, the NDVI needs to be aggregated at this scale. Doing so could lead to an "aggregate-out" problem. That is, if the treatment has a spatially localized impact, averaging the NDVI over the whole woreda could dilute its effect and lead us to misconclude to an absence of effect of the PSNP. A similar concern arises regarding the time dimension. Indeed, because effects of the PSNP should be concentrated over the growing season, averaging the NDVI over the full year could dilute the effect of the program.

To tackle these issues, we impose spatial and time constraints when aggregating the NDVI. Regarding the spatial dimension, we average the NDVI using pixels covering cultivated areas only. To do so, we rely on the Land Use database provided by MODIS (Friedl et al., 2010). This database, available on an annual basis, provides information on soil occupation (forest, savannas, grasslands, croplands, etc.) at a resolution of 500m. One may worry that land occupation could itself be affected by the program, as the PSNP may lead farmers to cultivate new plots (extensive margin effect). In that case, using yearly data on soil occupation to compute the NDVI could also lead to a downward bias in our estimates if pixels newly identified as cultivated areas have a structurally lower NDVI than older cultivated areas. For this reason, we

[13] We restrict our sample to this period due to data availability of control variables.

[14] We believe that the woreda level is relevant to study the impact of PSNP works because public works have been integrated into woredas development plans and because the targeting of woredas has remained remarkably stable over time. Woreda administrators allocated beneficiary numbers to kebeles using various approaches (either including all kebeles or selecting some of them) and this allocation varied over time. No data are available on treatment allocation at the kebele-year level.

focus only on plots that were cultivated at the onset of the program.[15] We compute the NDVI using pixels covering cultivated areas in 2005. Regarding the time dimension, we aggregate the NDVI over months covering the growing season of the main crop in each woreda. To do so, we use the MIRCA 2000 database (Portmann et al., 2010) which provides information on the type of crops, the area under cultivation, and the period of the growing season at a grid resolution of 5 arc-minute (10km).

In sum, for each woreda and each year of the 2000-2013 period, we compute the average NDVI using pixels covering cultivated areas in 2005, and months corresponding to the growing season of the main crop cultivated.

In order to check the validity of this indicator, we use the 2013 and 2015 LSMS-ISA survey rounds and investigate whether it is indeed a good predictor of crop production and crop productivity.[16] We derive both the total production and the average productivity of land in 2013 and 2015, and test whether these measures are well correlated with our indicator. The results support the idea that our indicator is a good predictor of agricultural output (Table 1). The indicator is positively and significantly correlated with both measures of agricultural output. Importantly, these relationships hold when woreda fixed effects are included, meaning that the indicator not only predicts levels of agricultural outputs (columns 1 and 4), but also their variations over time (columns 2 and 5). Overall, our indicator seems well suited to capture PSNP work effects at the intensive margin, i.e., productivity gains on parcels already cultivated when the program was launched in 2005.[17]

[15]We tried to design a satellite-based outcome to capture effects at the extensive margin but were admittedly unsuccessful (see footnote 17 for more information).

[16]An additional LSMS-ISA survey round was implemented in 2011. However, data on crop production are missing for most of the plots because of implementation issues. We are therefore unable to incorporate these data in our investigation of the crop productivity indicator.

[17]PSNP works may also have had effects at the extensive margin through land rehabilitation measures. We tried to design a satellite-based indicator to capture these effects but were admittedly unsuccessful. Specifically, for each year of the 2001-2013 period, we derived the share of cultivated areas by woredas using MODIS Land Use database, and checked the predictive power of this variable using cultivated areas derived from the 2013 and 2015 LSMS-ISA surveys. While satellite-based cultivated area appear to be a good predictor in levels, it does not appear to be capturing variations properly (Table A2). This could be due to a high level of inertia in the data or to the fact that remote sensed data are not error-free (Gibson et al., 2020; Gibson, 2020), especially in the case of binary variables (Alix-Garcia and Millimet, 2020). This signals that the use of this proxy is not suited for our purpose and we therefore prefer to leave investigations on the extensive margin to future research.

Table 1: Correlation between satellite-based and survey-based agricultural production

	Production			Productivity		
	(1)	(2)	(3)	(4)	(5)	(6)
NDVI	4.825***	4.824**	6.806***	3.445***	4.837***	7.088***
	(1.043)	(1.884)	(2.071)	(0.822)	(1.734)	(1.823)
Woredas FE		✓	✓		✓	✓
Time FE			✓			✓
Observations	482	482	482	482	482	482
R-squared	0.10	0.83	0.85	0.08	0.72	0.76

Notes: Data on agricultural output comes from the Ethiopian 2013 and 2015 LSMS-ISA surveys. In columns (1)-(3), the dependent variable corresponds to the overall production in woreda w at time t (with $t = 2013 \mid 2015$). In columns (4)-(6), the dependent variable corresponds to the average production per hectare in woreda w at time t. An inverse hyperbolic sine (IHS) transformation has been applied to all dependent variables. OLS estimator used for all regressions. Standard errors in parentheses are clustered at woreda level. *** p<0.01, ** p<0.05, * p<0.1.

While these results are encouraging, it is worth noting that the correlation between our satellite-based indicator and the true (unobserved) agricultural output may be even stronger than suggested in Table 1 due to some shortcomings in the LSMS-ISA measures of crop production. First, GPS coordinates in LSMS-ISA datasets are slightly modified (with a random noise of 0-10km) to preserve anonymity and as a result some plots may have been assigned to a neighboring woreda. Second, survey-based agricultural production may lack precision because LSMS-ISA surveys are not designed to be representative at the woreda level,[18] and because data on harvest and cultivated area are typically recalled with errors (Beegle et al., 2012; Carletto et al., 2015). The fact that our indicator of agricultural productivity is still highly significantly correlated with these survey-based measures is reassuring. If anything, the true (unobserved) predictive power of our satellite-based indicator should be magnified.

3.2 Treatment variables

Data on program implementation are drawn from the United Nations Office for the Coordination of Humanitarian Affairs (UNOCHA). To investigate the impact of PSNP works, we use three definitions of the treatment. We first use a basic dummy variable taking the value one if a woreda receives the program and zero otherwise. Figure 1a represents the woredas that received the PSNP. Second, we use a treatment intensity variable, defined as the percentage of population targeted by the program (Figure 1b). Data on treatment intensity present two shortcomings. First, they provide intervals of treatment intensity instead of precise percentages. The intervals are the following: (i) 2-13%; (ii) 14-25%; (iii) 26-42%; (iv) 43-65%; and (v)

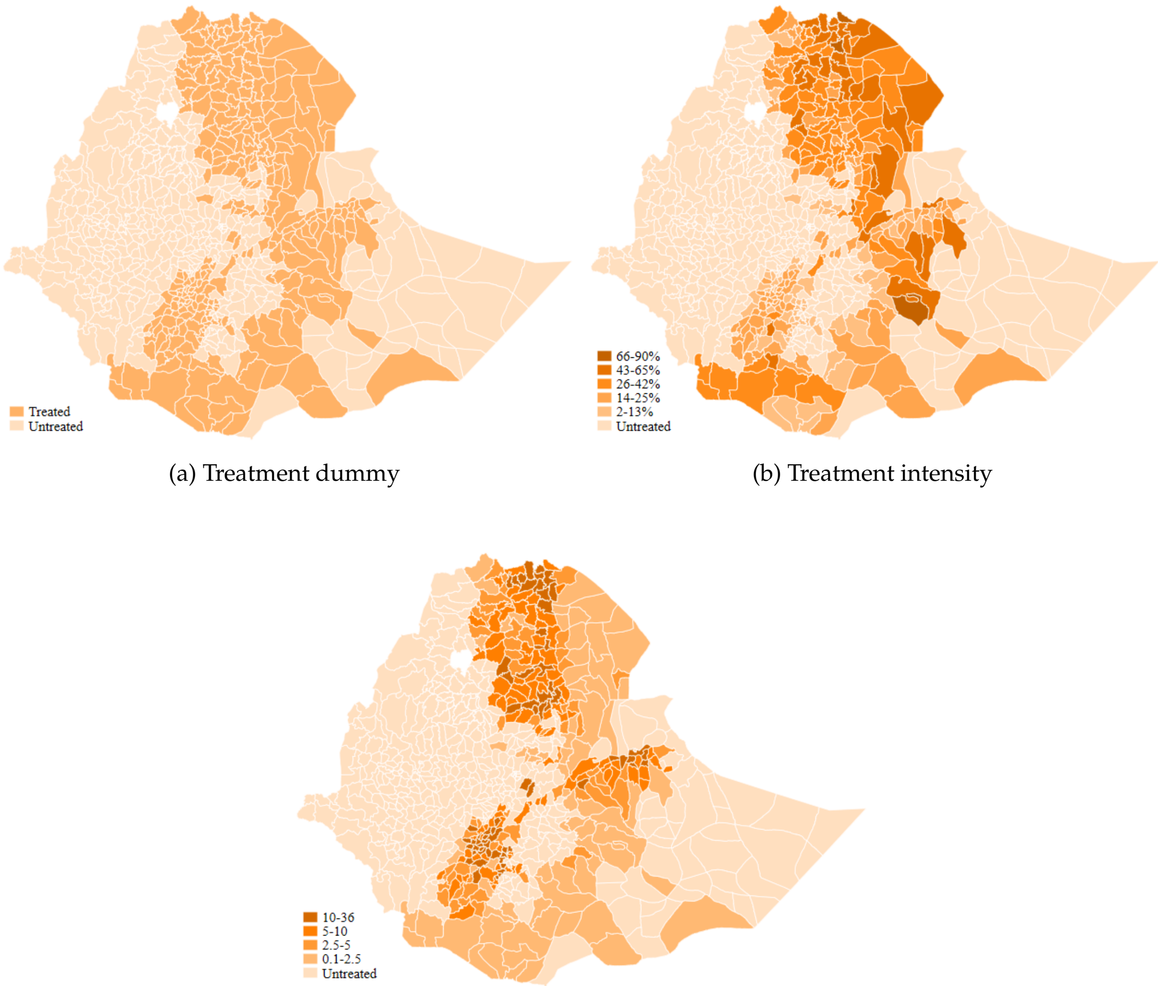
[18] In each enumeration area, 15 households were surveyed. Most woredas have only one enumeration area.

66-90%. We choose to use an ordered categorical variable taking the value 0 (if the woreda is not treated) to 5 (if the woreda received the highest treatment intensity). Second, we only have data on treatment intensity at the onset of the project in 2006.[19] Finally, for some woredas (mostly in the Somali region), we have no data on treatment intensity because they had not yet been assigned to a treatment intensity in 2006 (they joined the program much later). One possibility would be to drop these observations from the analysis. Alternatively, we could assign the average treatment intensity derived from the sample of woredas with positive values. To limit power losses, and because it concerns few woredas, we prefer the latter option. Last, we use a treatment density variable, defined as the number of beneficiaries per square kilometers in each woreda (Figure 1c). In particular, we rely on population density estimates from 2005, and approximate exact treatment intensity using the median value over each of the five interval (e.g. for the interval 2-13% we derive a treatment intensity of 7.5%).[20]

[19] However, both the intensity of treatment across regions and the allocation of treatment across woredas have remained stable over time (World Bank, 2016), suggesting that variations in the treatment intensity across woredas may also have remained stable.

[20] In subsequent analysis, we divide this variable by 10 to ease the reporting of estimated coefficients.

Figure 1: Woredas covered by the PSNP



(a) Treatment dummy

(b) Treatment intensity

(c) Treatment density

Notes: Figure 1a represents treated and untreated woredas; Figure 1b represents the percentage of beneficiaries by woreda; Figure 1c represents the number of beneficiaries per square kilometers. Source: Authors' elaboration from UNOCHA data.

4 Empirical analysis

In order to investigate the impact of PSNP works on crop production, we estimate the following difference-in-differences (DID) model using an ordinary least squares (OLS) estimator:

$$y_{wt} = \beta_0 + \beta_1 Treated_w \times Post_t + \mathbf{X}^{'} \alpha + \nu_w + \gamma_t + \varepsilon_{wt} \quad (1)$$

where β_1 gives the average treatment effect of interest; y_{wt} is the indicator of agricultural productivity for woreda w at time t; *Treated* corresponds to one of the three treatment variables defined in Section 3 (i.e. the treatment dummy, intensity, or density); *Post* is a dummy variable taking the value one for post-program years (2007-2013 in most woredas), and zero otherwise;[21] $\boldsymbol{X}$ is a vector of time varying control variables including rainfall, temperature, and their respective quadratic terms drawn from CHIRPS database;[22] ν_w is a vector of woreda fixed-effects controlling for time-invariant characteristics; γ_t is a vector of year fixed-effects controlling for common shocks; and ε_{wt} is the error term.

As mentioned above, Ethiopia is known for its large ecological disparities, especially between highlands and lowlands. Because these heterogeneities could be an important factor mediating PSNP effects, we conduct the analysis separately for highlands and lowlands. We follow Hurni (1983) and define as highlands all woredas with mean elevation higher than 1500m. As a robustness check, we present the main results by elevation deciles instead of using an arbitrary cut-off.

The crucial assumption underlying DID models is the parallel trends assumption, that is, in the absence of treatment, the difference between the treated and controls would have remained constant over time. This assumption seems rather strong in our setting because the program was targeted towards chronically food insecure woredas, i.e., woredas that required frequent food assistance prior to 2005 and could therefore present unobserved time-varying factors. One of the main advantage of our dataset is that it includes multiple time periods prior to PSNP roll-out so that we can actually check whether the parallel trends assumption holds prior to the program. Figures 2a and 2d plot the evolution of the outcome for treated and controls. While both figures are purely descriptive in nature, they tend to suggest that crop productivity did not follow similar paths in the two groups before the program. To test in a more systematic and comprehensive way whether there were specific trends prior to the treatment, we use a regression model similar to equation (1), but incorporating interactions between the treatment and each of the pre-program year dummy. Results are presented in Table 2. The significant interactions in columns 1-2 and 5-6 confirm the intuitions from Figures 2a and 2d. The two groups were already following distinct paths prior to public works implementation, and assuming parallel trends post-program would be a strong assumption.

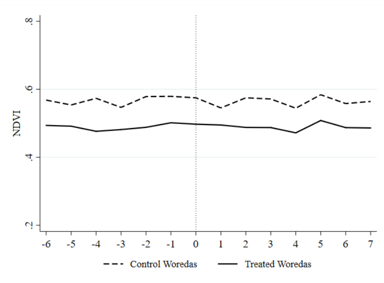
[21] For some treated woredas who started to receive public works only in 2009-10 the variable *Post* takes the value one only from 2010 onwards.

[22] We may be tempted to include additional variables such as night time lights or population density to control for economic and demographic dynamics. However, because these variables could be themselves affected by the treatment, we prefer to exclude them from the main model. In robustness analysis, we check whether including these variables affect the main results.

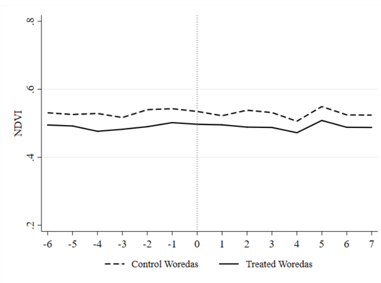
Figure 2: Agricultural productivity in treatment and control woredas

Panel A: Highlands

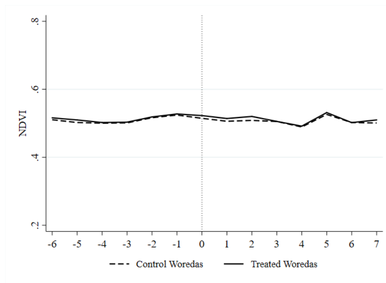
(a) Raw means



(b) Raw means
(common support)

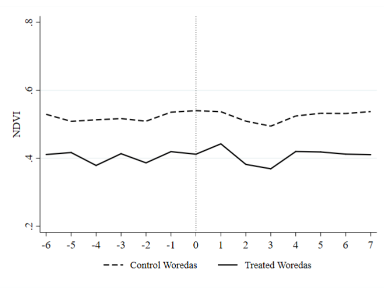


(c) IP-weighted means
(common support)

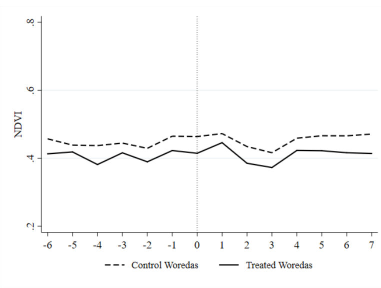


Panel B: Lowlands

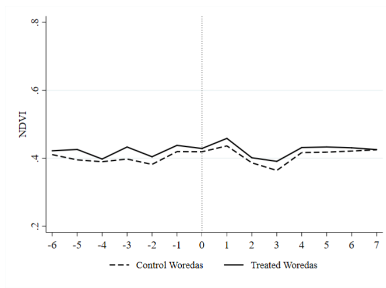
(d) Raw means



(e) Raw means
(common support)



(f) IP-weighted means
(common support)



Notes: Each sub-figure compares our indicator of agricultural productivity in treated and control woredas over time. Dotted lines display PSNP roll-out. Source: Authors' elaboration from MODIS and MIRCA 2000 data.

Table 2: Pre-treatment trends

	Highlands				Lowlands			
	(1) NDVI	(2) NDVI	(3) NDVI	(4) NDVI	(5) NDVI	(6) NDVI	(7) NDVI	(8) NDVI
Treatment × 2001	0.012***	0.014***	0.005	-0.001	0.027***	0.031***	0.027***	0.019***
	(0.001)	(0.003)	(0.004)	(0.004)	(0.003)	(0.005)	(0.006)	(0.007)
Treatment × 2002	-0.023***	-0.002	-0.001	-0.007	-0.016***	-0.001	0.003	0.000
	(0.002)	(0.003)	(0.004)	(0.005)	(0.003)	(0.005)	(0.006)	(0.006)
Treatment × 2003	0.009***	0.005	-0.001	-0.001	0.015***	0.003	0.001	0.006
	(0.002)	(0.003)	(0.004)	(0.004)	(0.003)	(0.005)	(0.006)	(0.006)
Treatment × 2004	-0.016***	-0.002	-0.001	0.004	-0.004*	0.007	0.013**	0.016**
	(0.002)	(0.003)	(0.004)	(0.004)	(0.002)	(0.005)	(0.006)	(0.006)
Treatment × 2005	-0.003**	-0.005*	-0.005	0.001	0.002	-0.002	-0.002	0.001
	(0.001)	(0.003)	(0.003)	(0.005)	(0.004)	(0.005)	(0.006)	(0.007)
Treatment × 2006	-0.003***	-0.003	-0.001	0.006	-0.010***	-0.004	0.002	0.003
	(0.001)	(0.003)	(0.003)	(0.005)	(0.003)	(0.005)	(0.005)	(0.006)
Woredas FE	✓	✓	✓	✓	✓	✓	✓	✓
Time FE	✓	✓	✓	✓	✓	✓	✓	✓
Time-varying controls		✓	✓	✓		✓	✓	✓
Common support			✓	✓			✓	✓
IP-weights				✓				✓
Observations	6594	6594	4606	4606	2786	2786	2268	2268
R-squared	0.93	0.94	0.92	0.92	0.96	0.97	0.96	0.96

Notes: This table tests for the presence of specific pre-program trends in agricultural productivity between treated and control woredas. The outcome variable is observed at the woreda-year level. Time varying controls include climatic variables (i.e. rainfall, temperature, and their respective quadratic terms). OLS estimator is used for all regressions, except regressions (4) and (8) where a WLS estimator with IP-weights is used. Only the interaction terms are reported due to space limitation. Standard errors in parentheses are clustered at the level of the treatment (*woredas*). *** p<0.01, ** p<0.05, * p<0.1.

One potential avenue to enrich our baseline DID model and remove these specific trends is to employ the Inverse Probability Weighting (IPW) method. This method, first pioneered by Horvitz and Thompson (1952), has been widely used in recent years to recover unbiased estimates of average treatment effects in observational studies (see e.g. Bernard et al., 2008; Austin and Stuart, 2015; Misha et al., 2019; Bargain et al., 2019).[23] We first estimate the following equation using a logit estimator:

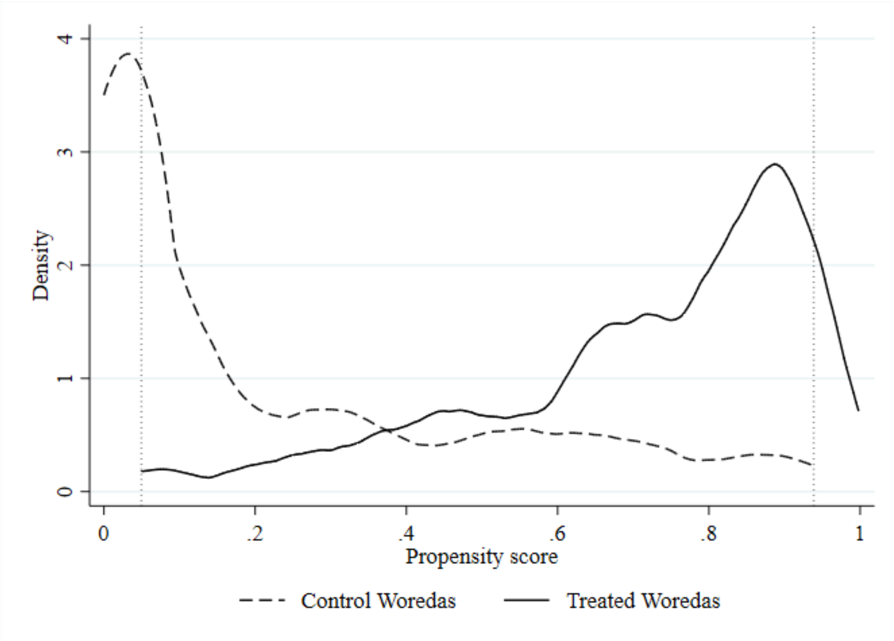
$$Treated_w = \alpha + \boldsymbol{X}_w^{'}\beta + \varepsilon_w \tag{2}$$

Where $\boldsymbol{X}_w$ is a vector of baseline covariates including climatic, geological, agricultural, demographic and economic determinants of treatment assignment. More specifically, $\boldsymbol{X}$ includes rainfall, temperature, elevation, slope, start and end months of the growing season, total area,

[23]The IPW method consists in computing the propensity score $e_w = Pr(T_w = 1 \mid X_w)$, where T_w is a dummy variable indicating whether woreda w is treated, and X_w is a vector of observed baseline covariates. Weights P_w are then derived as the inverse of the probability of receiving the treatment actually received.

share of cultivated area, population density, nighttime lights, and NDVI.[24] Each of these variables is averaged by woreda over the whole pre-treatment period. Then, we use estimates from equation (2) to predict propensity scores and derive weights P_w as the inverse probability of receiving the treatment actually received, and use them in a Weighted Least Squares (WLS) estimate equation (1).

Figure 3: Propensity scores distribution by treatment groups



Notes: This figure reports the distribution of the treatment assignment probabilities derived from equation (2) among treatment and control groups. Dotted lines display limits of the common support region. Source: Authors' elaboration.

The distribution of propensity scores in treated and control woredas are very different (Figure 3).[25] In particular, there are large spikes of (i) control woredas with low probabilities of treatment, and of (ii) treated woredas with high probabilities of treatment. These spikes suggest that the model specified in equation (2) is relatively successful at predicting assignment to treatment, and that using IPW techniques to estimate equation (1) has the potential to improve estimates. However, as can also be seen from the figure, there is a non-negligible share of woredas falling outside of the common support region. To avoid that these woredas affect

[24]Rainfall and temperature are drawn from CHIRPS database, elevation and slope are from the Shuttle Radar Topography Mission dataset (v4.1), start and end months of the growing season are from the MIRCA 2000 database, total area and share of cultivated area are from the MODIS land use database, population density is from Gridded Population of the World v4, nighttime lights is from DMSP-OLS (v4).

[25]Estimates of equation (2) are shown in Table A3.

our estimates, we exclude them from the main regressions.[26] In Figure A2, we show a map of propensity scores for woredas within the common support.

We test the validity of our common support restriction and IPW procedure in a variety of ways. First, by checking graphically whether differences in pre-program trends between treated and controls are attenuated (Figure 2). Restricting the sample to the common support region reduces the gap between the two curves and seems to slightly improve their parallelism (sub-figures 2b and 2e). IPW procedures further reduce differences between the two curves and pre-program trends become relatively difficult to distinguish (sub-figures 2c and 2f). Second, we test more formally for the existence of significant differences in pre-trends (Table 2). Results confirm that the procedures described above successfully remove specific pre-trends. Both the common support restriction and the use of IPW reduce the magnitude and significance of interaction terms, especially for highland woredas where no interaction terms are statistically significant at conventional levels (column 4).[27] Finally, we conduct balance tests on the common support sample. Results clearly indicate that our procedures allow to balance pre-program characteristics in the two groups (Table 3). Most of the 11 characteristics tested show significant differences using unweighted means, whereas none of these differences are significant using IP-weighted means. Most importantly, the F-test of the omnibus test for joint significance is very low and non-significant in the IPW case. Overall, these tests provide some reassurance on the validity of our procedures and on our ability to recover credible estimates of public works effects. The next section presents the main results.

[26]IPW techniques typically allocate low weights to woredas outside the common support region. However, because of the relatively large number of woredas concerned in our setting, we prefer to go one step further and exclude these woredas to prevent them from having any influence on the estimates. In robustness analysis, we replicate our main analysis keeping these woredas.

[27]Treated woredas in lowlands experienced a significant improvement of their productivity in 2001 and 2004 that is not captured by our model. However, this improvement is not too worrying because it is relatively small in magnitude and does not persist over time.

Table 3: Pre-treatment characteristics by sub-samples

	Raw means			IP-weighted means		
	(1) Treated	(2) Controls	(3) Diff	(4) Treated	(5) Controls	(6) Diff
Propensity score	0.69	0.36	0.33***	0.51	0.57	-0.06
			(0.02)			(0.05)
NDVI	0.46	0.51	-0.04***	0.48	0.48	0.00
			(0.01)			(0.01)
Rainfall	106.46	119.33	-12.86***	119.05	115.12	3.93
			(4.58)			(7.00)
Temperature	29.51	27.52	2.00***	28.40	28.62	-0.22
			(0.53)			(0.66)
Total area	0.14	0.14	0.00	0.14	0.14	0.00
			(0.02)			(0.02)
Cultivated area (% total area)	0.21	0.17	0.04**	0.23	0.23	0.00
			(0.02)			(0.03)
Start growing season	6.02	6.06	-0.04	6.03	6.04	-0.01
			(0.08)			(0.06)
End growing season	10.03	9.82	0.21**	9.99	10.00	-0.01
			(0.10)			(0.09)
Elevation	1689.61	1759.49	-69.88	1791.37	1734.06	57.31
			(59.12)			(82.05)
Slope	5.46	4.56	0.90***	4.98	5.55	-0.57
			(0.28)			(0.53)
Population density	31.05	31.61	-0.55	27.68	30.31	-2.63
			(6.99)			(4.43)
Night time lights	0.10	0.21	-0.12	0.09	0.13	-0.05
			(0.14)			(0.08)
Observations	266	225	491	266	225	491
F-test joint significance			24.07***			0.52

Notes: Sample trimmed to common support region. The F-test corresponds to a regression of the treatment on baseline characteristics using the same specification as in subsequent analysis (omnibus test). Standard errors in parentheses are clustered at the level of the treatment (*woredas*). *** p<0.01, ** p<0.05, * p<0.1.

5 Results

5.1 Treatment effects on crop productivity

Throughout the specifications, we find no evidence to suggest that public works increased agricultural productivity in beneficiary woredas. Estimates of equation (1) using IPW and the common support restriction are reported in Table 4. Columns 1-3 report the results for each of the treatment variable on the sample of highland woredas. Columns 4-6 report results on the sample of lowland woredas. All estimates control for woreda fixed effects, time fixed effects, and time-varying controls including climatic variables, i.e., rainfall, temperature and their respective quadratic terms. Point estimates suggest that benefiting from public works had small and non-significant effects in both highlands and lowlands. These results hold regardless of the definition of the treatment variable, suggesting that neither the least treated nor the least densely populated woredas drive these null estimates.

Table 4: Impacts on agricultural output

	Highlands			Lowlands		
	NDVI	NDVI	NDVI	NDVI	NDVI	NDVI
	(1)	(2)	(3)	(4)	(5)	(6)
Post × Treatment (dummy)	0.001			-0.002		
	(0.002)			(0.003)		
Post × Treatment (intensity)		0.000			-0.001	
		(0.001)			(0.001)	
Post × Treatment (density)			0.001			0.000
			(0.001)			(0.003)
Woredas FE	✓	✓	✓	✓	✓	✓
Time FE	✓	✓	✓	✓	✓	✓
Time-varying controls	✓	✓	✓	✓	✓	✓
IP-weights	✓	✓	✓	✓	✓	✓
Mean dependent variable	0.51	0.51	0.51	0.43	0.43	0.43
SD dependent variable	0.08	0.08	0.08	0.13	0.13	0.13
Observations	4606	4606	4606	2268	2268	2268
R-squared	0.91	0.91	0.91	0.96	0.96	0.96

Notes: Sample trimmed to common support region. Time varying controls include climatic variables (i.e. rainfall, temperature, and their respective quadratic terms). WLS estimator is used for all regressions. Standard errors in parentheses are clustered at the level of the treatment (*woredas*). *** p<0.01, ** p<0.05, * p<0.1.

One typical concern with non-significant results is that they can sometimes reflect a lack of statistical power rather than a lack of positive effects. We argue that this is not the case in this study for at least two reasons. First, estimates on the sample of lowland woredas actually present a negative sign – an attribute inconsistent with positive effects being obscured by insufficient power. Second, we can rule out even small positive effects for highland woredas. The upper bound of the 95 percent confidence interval on estimates for the treatment dummy is 0.005. Given the relationship of our NDVI indicator with the survey-based measure of agricultural productivity outlined in Table 1, this upper bound corresponds to a 3.6% increase in agricultural productivity. Following Ioannidis et al. (2017), we can derive the minimum detectable effect size (MDES) at conventional power (80%) and statistical significance (5%) by multiplying standard errors by 2.8. Using the highest standard errors of Table 4 and estimates from Table 1, we find that our study is powered to detect productivity gains above 6%.[28]

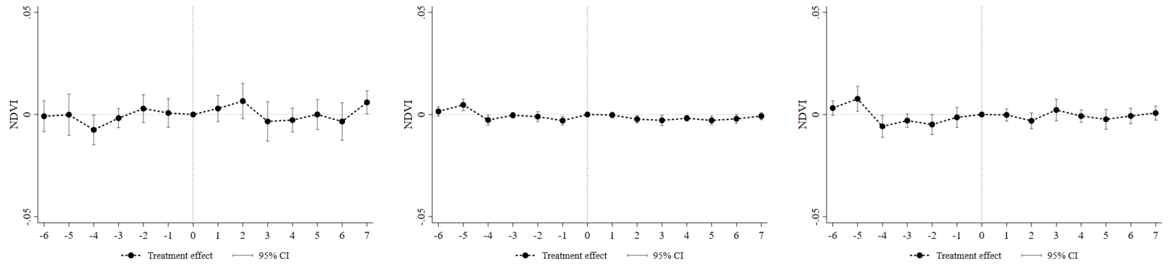
The evidence presented above is indicative of null effects over the whole 2005-2013 period. Nevertheless, null effects could mask subtle temporal patterns. In Figure 4, we explore the evolution of treatment effects over time. Because positive effects of public works could take time to manifest, and because public works received by beneficiary woredas naturally accumu-

[28]Magnitude of impacts are often compared across studies using standard deviations. In Table A4, we replicate our results using a standardized outcome variable. We find point estimates of no more than 0.018 SD, with small standard errors. Using the highest standard errors of Table A4, we estimate an MDES of 0.07 SD. Such an effect is generally considered as small in the literature.

Figure 4: Treatment effects over time

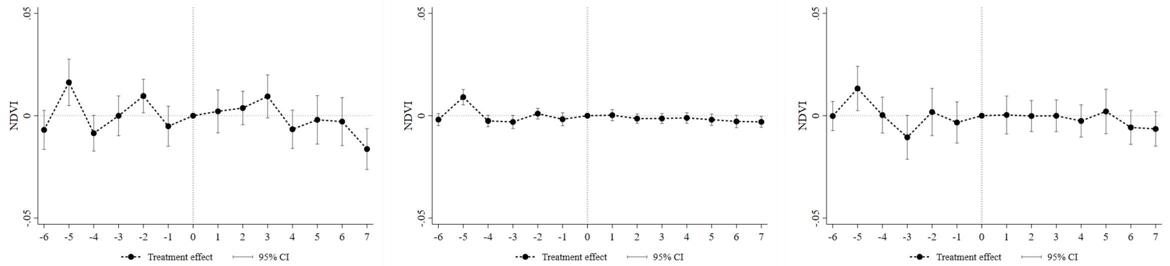
Panel A: Highlands

(a) Treatment (dummy) (b) Treatment (intensity) (c) Treatment (density)



Panel B: Lowlands

(d) Treatment (dummy) (e) Treatment (intensity) (f) Treatment (density)



Notes: Figures represent the evolution of treatment effects over time. Dotted vertical lines display PSNP roll-out.

late over time, we might expect to see a sustained increase in observed impacts over the course of the program. In practice, we find no evidence of such an increase. Treatment effects in late years are not particularly larger nor more significant than treatment effects in early years, and importantly there is no obvious upward trend in estimated treatment effects over the period considered in the analysis.

Finally, the impact of public works could be conditional on climatic conditions. In particular, the nature of PSNP works (e.g. land improvements, soil and water conservation measures) could help to mitigate adverse climate shocks such as droughts. For instance, a World Bank report argues that "*the works have been found to bring demonstrable benefits to farmers from the conservation of moisture, which not only leads to visibly improved plant growth close to the bunds, but also to an increase in ground water recharge such that dry springs have started to flow again and local stream flows have increased*" (World Bank, 2006). This improvement in water resources availability and management could make beneficiary woredas more resilient to rainfall deviations. To explore these potential effects, we incorporate a triple-interaction $Treated_w \times Post_t \times Rainfall_{wt}$

in our main model.[29] As can be seen from the signs of the triple-interactions, crop productivity in beneficiary woredas seems to be less sensitive to rainfall (Table 5), which would be consistent with the idea that the PSNP increased resilience to climatic conditions. However, evidence on these effects remain limited as only one of the six interactions is significant at conventional levels, which could reflect multiple testing issues.

Table 5: Triple difference

	Highlands			Lowlands		
	NDVI	NDVI	NDVI	NDVI	NDVI	NDVI
	(1)	(2)	(3)	(4)	(5)	(6)
Post × Treatment (dummy) × Rainfall	-0.005			-0.006		
	(0.005)			(0.005)		
Post × Treatment (dummy)	0.008			0.004		
	(0.007)			(0.005)		
Post × Treatment (intensity) × Rainfall		0.000			-0.005**	
		(0.002)			(0.002)	
Post × Treatment (intensity)		0.000			0.002	
		(0.002)			(0.002)	
Post × Treatment (density) × Rainfall			0.000			0.005
			(0.004)			(0.012)
Post × Treatment (density)			0.000			-0.003
			(0.005)			(0.010)
Woredas FE	✓	✓	✓	✓	✓	✓
Time FE	✓	✓	✓	✓	✓	✓
Time-varying controls	✓	✓	✓	✓	✓	✓
IP-weights	✓	✓	✓	✓	✓	✓
Mean dependent variable	0.51	0.51	0.51	0.43	0.43	0.43
SD dependent variable	0.08	0.08	0.08	0.13	0.13	0.13
Observations	4606	4606	4606	2268	2268	2268
R-squared	0.91	0.91	0.91	0.96	0.96	0.96

Notes: Annual rainfall expressed in meters. Standard errors in parentheses are clustered at the level of the treatment (*woredas*). *** p<0.01, ** p<0.05, * p<0.1. See notes to Table 4 for other details.

5.2 Robustness checks

We first explore the robustness of our findings to five variations to the main specification: (i) adding control variables to account for economic and demographic trends; (ii) using Oster bounds to establish the likely degree of omitted variable bias (Oster, 2019); (iii) estimating the main effects keeping woredas outside of the common support region; (iv) estimating the main effects keeping the most treated woredas; and (v) estimating the main effects by elevation deciles instead of using an ad hoc cut-off to distinguish lowlands and highlands. Results, presented respectively in Tables A5, A6, A7, A8, and Figure A3, are qualitatively unchanged. Adding control variables, keeping woredas outside of the common support region, or keeping woredas with higher treatment density, no coefficients become significant. Oster bounds

[29] We naturally also include all other relevant interactions.

suggest that residual omitted variable bias is very small and unlikely to explain lack of effects (Table A6). No clear pattern is visible from estimates by elevation deciles.

Then, we investigate alternative explanations for the observed effects. Two main stories could threaten our interpretation: (i) a negative effect of PSNP transfers on agricultural productivity; (ii) an increase in net emigration from beneficiary woredas. Regarding the first threat, as mentioned in Section 2, beneficiary households received cash or food transfers in exchange for their participation in public works. This could be problematic for our estimates if these transfers had a negative effect on agricultural productivity, creating a downward bias in the estimates. We argue that it is unlikely to be the case for at least two reasons. First, the literature actually suggests that PSNP transfers had modest effects on crop productivity (Hoddinott et al., 2012). Second, we investigate whether PSNP transfers could have diverted beneficiaries from agriculture by testing whether the share of pixels cultivated in 2005 and still cultivated in 2013 are affected by the treatment. Table A9 suggests that the program had no such effects.

An increase in net emigration from beneficiary woredas could also introduce a downward bias in the estimates by reducing the availability of labor for agriculture. Theoretically, the impact on net emigration of a social protection program such as the PSNP is ambiguous. On the one hand, it could increase emigration by relaxing financial and risk constraints (Angelucci, 2015; Gazeaud et al., 2019a). On the other hand, it could reduce emigration through increased opportunity costs (Imbert and Papp, 2020), or increase immigration by making beneficiary woredas more attractive to aspiring migrants. Because of a lack of data on migration flows, especially at relatively disaggregated levels, it is empirically challenging to investigate program effects on net emigration. Using panel data covering the 2006-2012 period, Hoddinott and Mekasha (2020) find no evidence suggesting that participation in the PSNP led to an increase in emigration. In fact, the authors find that the program significantly decreased emigration of adolescent girls. Using data from the 2007 census, we provide suggestive evidence that the program did not increase net emigration from beneficiary woredas in the early years of the program.[30] We first compute domestic immigration rates (per 1,000 individuals) for each woreda over the 2000-2007 period, and then check whether the program had any effect using the main specification from Section 4. As can be seen from Table A10, the program did not seem to impact significantly immigration to beneficiary woredas. Because domestic immigration and domestic emigration are two sides of the same coin, and international migration flows are negligible in rural Ethiopia,[31] we argue that measuring the effect of the program on domes-

[30] A census was conducted in 2017 but data still have to be released.

[31] For example, in 2007, international immigrants represent only 0.1% of all Ethiopians and 0.8% of all immigrants (authors' estimates using data from the 2007 census).

tic immigration is actually the reverse of measuring the effect of the program on net emigration. The lack of impact on agricultural productivity does not appear to be explained by an increase in net emigration from beneficiary woredas.

6 Conclusions

African countries are particularly vulnerable to climate change (Kurukulasuriya et al., 2006) and farm households in these countries have notoriously few options to cope with climate shocks (Fafchamps et al., 1998). Therefore, it leaves important responsibilities to policy makers to build appropriate policies. In 2005, the Ethiopian government launched an unusually large and durable public works program, called the PSNP, in the aim of mitigating the effects of climate shocks and promoting long-term development through increased productivity. The present paper provides novel evidence on the potential value of the infrastructure built under the program. We rely on an original dataset covering whole Ethiopia over the 2000-2013 period, and explore the effect of the program using difference-in-differences estimates in combination with the inverse probability weighting method. Our result is a disappointing precise zero, meaning that there is no discernible effects of the PSNP on crop productivity.

To test the validity of this result, we run several robustness checks. First, as Ethiopia has important ecological disparities across its territory, we run separate regressions for highlands and lowlands, using different thresholds. Estimates remain non significant and close to zero, suggesting that null effects do not hide heterogeneous impacts along topographical characteristics. Second, to check that our result is not driven by a lack of statistical power, we compute the minimum detectable effect size and provide evidence that we would be able to detect even small effects. We also compute Oster bounds and show that residual omitted variables are unlikely to drive the lack of effect. Last, as the program could take time to produce benefits, we estimate the effect of the PSNP on a year by year basis. Again, we find no evidence of positive impacts even seven years after the beginning of the program. Finally, regarding the potential transmission channels, we provide suggestive evidence that neither internal migration nor labor reallocation patterns explain the null effects.

Nevertheless, our results should in no way be interpreted as definitive evidence against PWP. We focus here on the Ethiopian PSNP, a pioneering program which has inspired many similar programs in other African countries (Subbarao et al., 2012). While evidence on this program is of obvious interest, one should be cautious about generalizing from the PSNP to other contexts. In addition, other infrastructures were built under the PSNP and they may have brought benefits that do not necessarily show up on crop production, such as in the case of road investments which may have caused important time savings (World Bank, 2010). Finally, we believe that further research is needed to better understand the results highlighted in this paper. In particular, there is little we can say on the reasons that may explain the lack of effects on productivity. Potential explanations relate to the low quality or lack of durability of the infrastructure and to possible crowding-out effects on private investments in soil and water conservation measures. For instance, there is evidence outside the PSNP that infrastructures generated in PWP can be inappropriate, of poor quality, or not maintained (Barrett et al., 2004; Holden et al., 2006; Kaur et al., 2019; Gazeaud et al., 2019b), and that public works can crowd-out private investments in land conservation measures (Gebremedhin and Swinton, 2001; Barrett et al., 2004). Formally addressing this question in the context of the PSNP requires further data collection and is left for future research.

References

Adimassu, Z. and Kessler, A. (2015). Impact of the productive safety net program on farmers' investments in sustainable land management in the Central Rift Valley of Ethiopia. *Environmental Development*, 16:54–62.

Ajefu, J. B. and Abiona, O. (2019). Impact of shocks on labour and schooling outcomes and the role of public work programmes in rural India. *The Journal of Development Studies*, 55(6):1140–1157.

Alderman, H. and Yemtsov, R. (2014). How can safety nets contribute to economic growth? *The World Bank Economic Review*, 28(1):1–20.

Alem, Y. and Broussard, N. H. (2018). The impact of safety nets on technology adoption: a difference-in-differences analysis. *Agricultural economics*, 49(1):13–24.

Ali, D. A., Deininger, K., and Monchuk, D. (2020). Using satellite imagery to assess impacts of soil and water conservation measures: Evidence from Ethiopia's Tana-Beles watershed. *Ecological Economics*, 169:106512.

Alik-Lagrange, A. and Ravallion, M. (2015). Inconsistent policy evaluation: A case study for a large workfare program. NBER Working Paper no. 21041.

Alik-Lagrange, A. and Ravallion, M. (2018). Workfare versus transfers in rural India. *World Development*, 112:244–258.

Alix-Garcia, J. and Millimet, D. L. (2020). Remotely incorrect? Manuscript.

Andersson, C., Mekonnen, A., and Stage, J. (2011). Impacts of the productive safety net program in Ethiopia on livestock and tree holdings of rural households. *Journal of Development Economics*, 94(1):119–126.

Angelucci, M. (2015). Migration and financial constraints: Evidence from Mexico. *Review of Economics and Statistics*, 97(1):224–228.

Araya, G. B. (2020). Impact of Ethiopia's productive safety net program on manure use by rural households: Evidence from Tigrai, Northern Ethiopia. *Agricultural Economics*, 51(5):725–742.

Asher, S. and Novosad, P. (2020). *Rural roads and local economic development*. JSTOR.

Atzberger, C. (2013). Advances in remote sensing of agriculture: Context description, existing operational monitoring systems and major information needs. *Remote sensing*, 5(2):949–981.

Austin, P. C. and Stuart, E. A. (2015). Moving towards best practice when using inverse probability of treatment weighting (IPTW) using the propensity score to estimate causal treatment effects in observational studies. *Statistics in medicine*, 34(28):3661–3679.

Baird, S., Ferreira, F. H., Özler, B., and Woolcock, M. (2014). Conditional, unconditional and everything in between: A systematic review of the effects of cash transfer programmes on schooling outcomes. *Journal of Development Effectiveness*, 6(1):1–43.

Bargain, O., Boutin, D., and Champeaux, H. (2019). Women's political participation and intrahousehold empowerment: Evidence from the Egyptian arab spring. *Journal of Development Economics*, 141:102379.

Barrett, C. B., Holden, S., and Clay, D. (2004). Can food-for-work programmes reduce vulnerability? In Dercon, S., editor, *Insurance Against Poverty*, pages 361–386. UNU-WIDER Studies in Development Economics.

Beegle, K., Carletto, C., and Himelein, K. (2012). Reliability of recall in agricultural data. *Journal of Development Economics*, 1(98):34–41.

Beegle, K., Galasso, E., and Goldberg, J. (2017). Direct and indirect effects of Malawi's public works program on food security. *Journal of Development Economics*, 128:1–23.

BenYishay, A., Runfola, D., Trichler, R., Dolan, C., Goodman, S., Parks, B., Tanner, J., Heuser, S., Batra, G., and Anand, A. (2017). A primer on geospatial impact evaluation methods, tools, and applications. AidData Working Paper no. 44.

Berg, E., Bhattacharyya, S., Rajasekhar, D., and Manjula, R. (2018). Can public works increase equilibrium wages? Evidence from India's national rural employment guarantee. *World Development*, 103:239–254.

Berhane, G., Gilligan, D. O., Hoddinott, J., Kumar, N., and Taffesse, A. S. (2014). Can social protection work in Africa? The impact of Ethiopia's productive safety net programme. *Economic Development and Cultural Change*, 63(1):1–26.

Bernard, T., Taffesse, A. S., and Gabre-Madhin, E. (2008). Impact of cooperatives on smallholders' commercialization behavior: evidence from Ethiopia. *Agricultural Economics*, 39(2):147–161.

Bertrand, M., Crépon, B., Marguerie, A., and Premand, P. (2017). Contemporaneous and post-program impacts of a public works program: Evidence from Côte d'Ivoire. Working Paper.

Besley, T. and Coate, S. (1992). Workfare versus welfare: Incentive arguments for work requirements in poverty-alleviation programs. *American Economic Review*, 82(1):249–261.

Burke, M. and Lobell, D. B. (2017). Satellite-based assessment of yield variation and its determinants in smallholder African systems. *Proceedings of the National Academy of Sciences*, 114(9):2189–2194.

Carletto, C., Gourlay, S., and Winters, P. (2015). From guesstimates to GPStimates: Land area measurement and implications for agricultural analysis. *Journal of African Economies*, 24(5):593–628.

Christian, S., de Janvry, A., Egel, D., and Sadoulet, E. (2015). Quantitative evaluation of the social fund for development labor intensive works program (LIWP). Working Paper.

Debela, B. L., Shively, G., and Holden, S. T. (2015). Does Ethiopia's productive safety net program improve child nutrition? *Food Security*, 7(6):1273–1289.

Deininger, K., Nagarajan, H. K., and Singh, S. K. (2016). *Short-term effects of India's employment guarantee program on labor markets and agricultural productivity*. Washington, DC: World Bank.

Devereux, S. (2000). Famine in the twentieth century. IDS Working Paper no. 105.

Dutta, P., Murgai, R., Ravallion, M., and Van de Walle, D. (2014). *Right to work?: Assessing India's employment guarantee scheme in Bihar*. The World Bank.

Fafchamps, M., Udry, C., and Czukas, K. (1998). Drought and saving in West Africa: are livestock a buffer stock? *Journal of Development economics*, 55(2):273–305.

Favara, M., Porter, C., and Woldehanna, T. (2019). Smarter through social protection? Evaluating the impact of Ethiopia's safety-net on child cognitive abilities. *Oxford Development Studies*, 47(1):79–96.

Filipski, M., Taylor, J. E., Abegaz, G. A., Ferede, T., Taffesse, A. S., and Diao, X. (2016). General equilibrium impact assessment of the productive safety net program in Ethiopia, 3ie Grantee Final Report. *International Initiative for Impact Evaluation (3ie)*.

Friedl, M. A., Sulla-Menashe, D., Tan, B., Schneider, A., Ramankutty, N., Sibley, A., and Huang, X. (2010). Modis collection 5 global land cover: Algorithm refinements and characterization of new datasets. *Remote sensing of Environment*, 114(1):168–182.

Galasso, E. and Ravallion, M. (2004). Social protection in a crisis: Argentina's plan jefes y jefas. *The World Bank Economic Review*, 18(3):367–399.

Gazeaud, J., Mvukiyehe, E., and Sterck, O. (2019a). Cash transfers and migration: Theory and evidence from a randomized controlled trial. CSAE Working Paper WPS/2019-16.

Gazeaud, J., Mvukiyehe, E., and Sterck, O. (2019b). Public works and welfare: A randomized control trial of the Comoros social safety net project. Endline Report.

Gebremedhin, B. and Swinton, S. M. (2001). Reconciling food-for-work project feasibility with food aid targeting in Tigray, Ethiopia. *Food Policy*, 26(1):85–95.

Gehrke, E. (2019). An employment guarantee as risk insurance? Assessing the effects of the NREGS on agricultural production decisions. *The World Bank Economic Review*, 33(2):413–435.

Gehrke, E. and Hartwig, R. (2018). Productive effects of public works programs: What do we know? What should we know? *World Development*, 107:111–124.

Gelman, A. and Carlin, J. (2014). Beyond power calculations: Assessing type s (sign) and type m (magnitude) errors. *Perspectives on Psychological Science*, 9(6):641–651.

Gibson, J. (2020). Better night lights data, for longer. CSAE Working Paper WPS/2020-08.

Gibson, J., Olivia, S., and Boe-Gibson, G. (2020). Night lights in economics: Sources and uses. Etudes et Documents, CERDI.

Gilligan, D. O., Hoddinott, J., and Taffesse, A. S. (2009). The impact of Ethiopia's productive safety net programme and its linkages. *Journal of Development Studies*, 45(10):1684–1706.

Haushofer, J. and Shapiro, J. (2016). The short-term impact of unconditional cash transfers to the poor: Experimental evidence from Kenya. *The Quarterly Journal of Economics*, 131(4):1973–2042.

Hirano, K., Imbens, G. W., and Ridder, G. (2003). Efficient estimation of average treatment effects using the estimated propensity score. *Econometrica*, 71(4):1161–1189.

Hirvonen, K. and Hoddinott, J. (2020). Beneficiary views on cash and in-kind payments: Evidence from Ethiopia's productive safety net programme. *The World Bank Economic Review (forthcoming)*.

Hoddinott, J., Berhane, G., Gilligan, D. O., Kumar, N., and Seyoum Taffesse, A. (2012). The impact of Ethiopia's productive safety net programme and related transfers on agricultural productivity. *Journal of African Economies*, 21(5):761–786.

Hoddinott, J., Gilligan, D. O., and Taffesse, A. S. (2010). The impact of Ethiopia's productive safety net program on schooling and child labor. *Social protection for Africa's children*.

Hoddinott, J. and Mekasha, T. J. (2020). Social protection, household size, and its determinants: Evidence from Ethiopia. *The Journal of Development Studies*, pages 1–20.

Holden, S., Barrett, C. B., and Hagos, F. (2006). Food-for-work for poverty reduction and the promotion of sustainable land use: can it work? *Environment and Development Economics*, 18:15–38.

Horvitz, D. and Thompson, D. (1952). A generalization of sampling without replacement from a finite population. *Journal of the American Statistical Association*, 47:663–685.

Hurni, H. (1983). Ethiopian highlands reclamation study. Food and Agriculture Organization of the United Nations (FAO).

Imbert, C. and Papp, J. (2015). Labor market effects of social programs: Evidence from India's employment guarantee. *American Economic Journal: Applied Economics*, 7(2):233–63.

Imbert, C. and Papp, J. (2020). Short-term migration, rural public works, and urban labor markets: Evidence from india. *Journal of the European Economic Association*, 18(2):927–963.

Ioannidis, J. P., Stanley, T., and Doucouliagos, H. (2017). The power of bias in economics research. *The Economic Journal*, 127(605):F236–F265.

Jayne, T. S., Strauss, J., Yamano, T., and Molla, D. (2002). Targeting of food aid in rural Ethiopia: Chronic need or inertia? *Journal of Development Economics*, 68(2):247–288.

Jensen, N., Stoeffler, Q., Fava, F., Vrieling, A., Atzberger, C., Meroni, M., Mude, A., and Carter, M. (2019). Does the design matter? Comparing satellite-based indices for insuring pastoralists against drought. *Ecological economics*, 162:59–73.

Kaur, N., Agrawal, A., Steinbach, D., Panjiyar, A., Saigal, S., Manuel, C., Barnwal, A., Shakya, C., Norton, A., Kumar, N., Soanes, M., and Venkataramani, V. (2019). Building resilience to climate change through social protection: Lessons from MGNREGS, India. IIED Working Paper.

Kehler, A. (2004). When will Ethiopia stop asking for food aid. *Humanitarion Exchange Magazine*, 27:1–4.

Klisch, A. and Atzberger, C. (2016). Operational drought monitoring in Kenya using MODIS NDVI time series. *Remote Sensing*, 8(4):267.

Kurukulasuriya, P., Mendelsohn, R., Hassan, R., Benhin, J., Deressa, T., Diop, M., Eid, H. M., Fosu, K. Y., Gbetibouo, G., Jain, S., et al. (2006). Will african agriculture survive climate change? *The World Bank Economic Review*, 20(3):367–388.

Lech, M., Uitto, J. I., Harten, S., Batra, G., and Anand, A. (2018). Improving international development evaluation through geospatial data and analysis. *International Journal of Geospatial and Environmental Research*, 5(2):3.

Li, T. and Sekhri, S. (2020). The unintended consequences of employment-based safety net programs. *The World Bank Economic Review*, 34(1):164–178.

Lobell, D. B., Azzari, G., Burke, M., Gourlay, S., Jin, Z., Kilic, T., and Murray, S. (2020). Eyes in the sky, boots on the ground: Assessing satellite-and ground-based approaches to crop yield measurement and analysis. *American Journal of Agricultural Economics*, 102(1):202–219.

Maity, B. (2020). Consumption and time-use effects of India's employment guarantee and women's participation. *Economic Development and Cultural Change (forthcoming)*.

Mendola, M. and Negasi, M. Y. (2019). Nutritional and schooling impact of a cash transfer program in Ethiopia: A retrospective analysis of childhood experience. Development Working Papers no. 451, Centro Studi Luca d'Agliano, University of Milano.

Merfeld, J. D. (2019). Spatially heterogeneous effects of a public works program. *Journal of Development Economics*, 136:151–167.

Misha, F. A., Raza, W. A., Ara, J., and Van de Poel, E. (2019). How far does a big push really push? Long-term effects of an asset transfer program on employment trajectories. *Economic Development and Cultural Change*, 68(1):41–62.

Murgai, R., Ravallion, M., and Van de Walle, D. (2015). Is workfare cost-effective against poverty in a poor labor-surplus economy? *The World Bank Economic Review*, 30(3):413–445.

Nega, F., Mathijs, E., Deckers, J., Haile, M., Nyssen, J., and Tollens, E. (2010). Rural poverty dynamics and impact of intervention programs upon chronic and transitory poverty in Northern Ethiopia. *African development review*, 22(1):92–114.

Oster, E. (2019). Unobservable selection and coefficient stability: Theory and evidence. *Journal of Business & Economic Statistics*, 37(2):187–204.

Pettorelli, N., Vik, J. O., Mysterud, A., Gaillard, J.-M., Tucker, C. J., and Stenseth, N. C. (2005). Using the satellite-derived NDVI to assess ecological responses to environmental change. *Trends in ecology & evolution*, 20(9):503–510.

Porter, C. and Goyal, R. (2016). Social protection for all ages? Impacts of Ethiopia's productive safety net program on child nutrition. *Social Science & Medicine*, 159:92–99.

Portmann, F. T., Siebert, S., and Döll, P. (2010). Mirca2000—global monthly irrigated and rainfed crop areas around the year 2000: A new high-resolution data set for agricultural and hydrological modeling. *Global biogeochemical cycles*, 24(1).

Quisumbing, A. R. and Yohannes, Y. (2005). How fair is workfare? gender, public works, and employment in rural ethiopia. World Bank Policy Research Working Paper no. 3492.

Rahmato, D. (2013). Food security and safety nets: Assessment and challenges. In Rahmato, D., Pankhurst, A., and van Uffelen, J.-G., editors, *Food security, safety nets and social protection in Ethiopia*, pages 111–144. Forum for Social Studies (FSS) Addis Ababa.

Ravallion, M. (2019). Guaranteed employment or guaranteed income? *World Development*, 115:209–221.

Ravallion, M., Galasso, E., Lazo, T., and Philipp, E. (2005). What can ex-participants reveal about a program's impact? *Journal of Human Resources*, 40(1):208–230.

Ravi, S. and Engler, M. (2015). Workfare as an effective way to fight poverty: The case of India's NREGS. *World Development*, 67:57–71.

Roodman, D. (2009). A note on the theme of too many instruments. *Oxford Bulletin of Economics and statistics*, 71(1):135–158.

Rosas, N. and Sabarwal, S. (2016). *Public works as a productive safety net in a post-conflict setting: Evidence from a randomized evaluation in Sierra Leone*. The World Bank.

Sabates-Wheeler, R. and Devereux, S. (2010). Cash transfers and high food prices: Explaining outcomes on Ethiopia's productive safety net programme. *Food Policy*, 35(4):274–285.

Schinaia, G. (2016). Labour market effects of social programmes: Evidence from Ethiopia's productive safety net. Master's thesis, University of Oxford.

Shah, M. and Steinberg, B. M. (2019). Workfare and human capital investment: Evidence from India. *Journal of Human Resources*, pages 1117–9201R2.

Sharp, K., Brown, T., and Teshome, A. (2006). *Targeting Ethiopia's productive safety net programme (PSNP)*. London and Bristol, UK: Overseas Development Institute and the IDL Group.

Shiferaw, B. and Holden, S. T. (1998). Resource degradation and adoption of land conservation technologies in the Ethiopian highlands: a case study in Andit Tid, North Shewa. *Agricultural Economics*, 18(3):233–247.

Subbarao, K., Del Ninno, C., Andrews, C., and Rodríguez-Alas, C. (2012). *Public works as a safety net: Design, evidence, and implementation*. Washington, DC: World Bank.

Wang, J., Rich, P. M., Price, K. P., and Kettle, W. D. (2005). Relations between NDVI, grassland production, and crop yield in the central great plains. *Geocarto International*, 20(3):5–11.

World Bank (2006). Ethiopia Second Productive Safety Nets Project. Project Apraisal Document. Washington, DC: World Bank.

World Bank (2010). Designing and implementing a rural safety net in a low income setting: Lessons learned from Ethiopia's productive safety net program 2005-2009. Washington, DC: World Bank.

World Bank (2013). Coping with change: How Ethiopia's PSNP and HABP are building resilience to climate change. World Bank Washington, DC.

World Bank (2016). Ethiopia Third Productive Safety Nets Project. Implementation completion and results report. Washington, DC: World Bank.

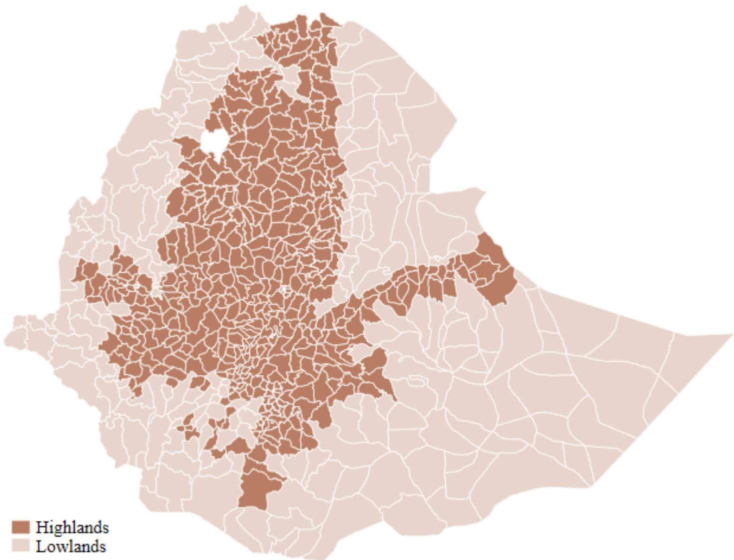
Zimmermann, L. (2012). Labor market impacts of a large-scale public works program: Evidence from the indian employment guarantee scheme. IZA Discussion paper no. 6858.

Table A1: Agricultural activities undertaken by Ethiopia's PSNP (2007-2009)

Project type	Output
Land rehabilitated through area closure	167,150 ha
Soil embankment construction	91,454 km
Stone embankment construction	184,730 km
Seedlings produced	1,321,938,020
Seedlings planted	883,321,700
Nursery sites established and managed	1,114
Ponds constructed or rehabilitated	133,927
Water springs developed	3,684
Hand-dug wells constructed	1,262
Small-scale irrigation canal control or rehabilitation	5,746 km

Source: Authors' elaboration from World Bank (2010) and Subbarao et al. (2012).

Figure A1: Map of highlands and lowlands



Notes: Highlands are defined as land with an altitude of 1500m or more.
Source: Authors' elaboration.

Table A2: Satellite and survey-based cultivated area

	(1) Cult. Area (LSMS-ISA)	(2) Cult. Area (LSMS-ISA)	(3) Cult. Area (LSMS-ISA)
Cult. Area (MODIS)	0.080***	-0.016	-0.014
	(0.022)	(0.017)	(0.018)
Woredas FE		✓	✓
Time FE			✓
Observations	482	482	482
R-squared	0.04	0.91	0.91

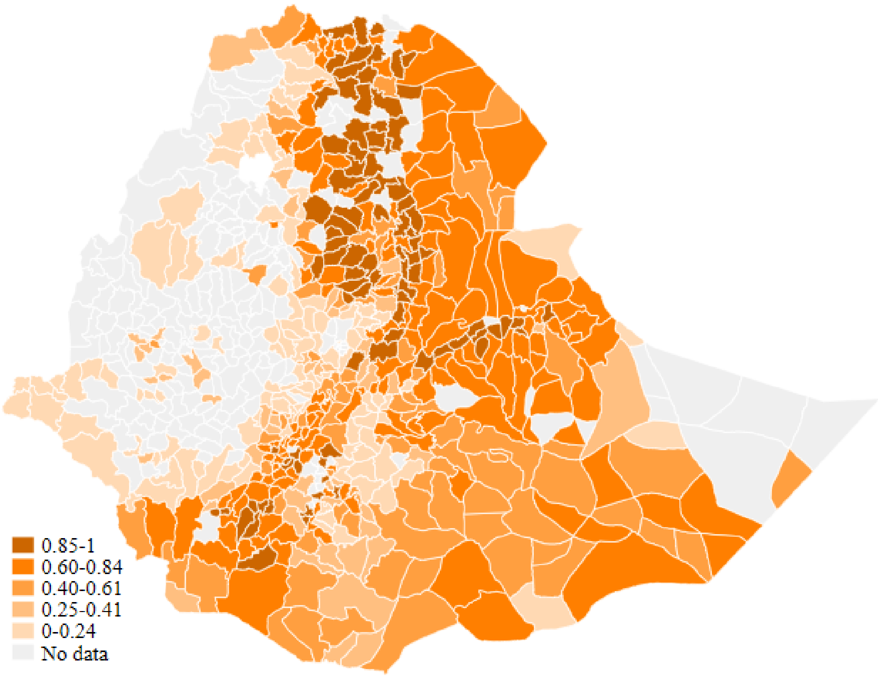
Notes: In columns (1)-(3), the dependant variable corresponds to the overall cultivated area in woreda w at time t (with $t = 2013 \mid 2015$), derived from the Ethiopian 2013 and 2015 LSMS-ISA surveys. An inverse hyperbolic sine transformation has been applied to all variables. Standard errors in parentheses are clustered at woreda level. *** p<0.01, ** p<0.05, * p<0.1.

Table A3: Determinants of the treatment

	(1) Treatment
Rainfall	0.060***
	(0.015)
Rainfall^2	-0.000***
	(0.000)
Temperature	0.712***
	(0.165)
Temperature^2	-0.009***
	(0.002)
Elevation	0.000
	(0.000)
Slope	0.289***
	(0.054)
Start growing season	0.695***
	(0.197)
End growing season	0.719***
	(0.209)
Total area	0.671
	(0.757)
Cultivated area (% total area)	0.583
	(1.043)
Population density	0.039***
	(0.006)
Night time lights	-1.862***
	(0.251)
NDVI	-7.173***
	(2.253)
Observations	670
R-squared	0.44

Notes: The outcome variable is a dummy equal to one if the woreda received the treatment. A logit estimator is used. Standard errors in parentheses are clustered at the woreda level. *** p<0.01, ** p<0.05, * p<0.1.

Figure A2: Map of propensity scores (common support)



Notes: Propensity scores are computed from Equation 2. Sample restricted to woredas on the common support.
Source: Authors' elaboration.

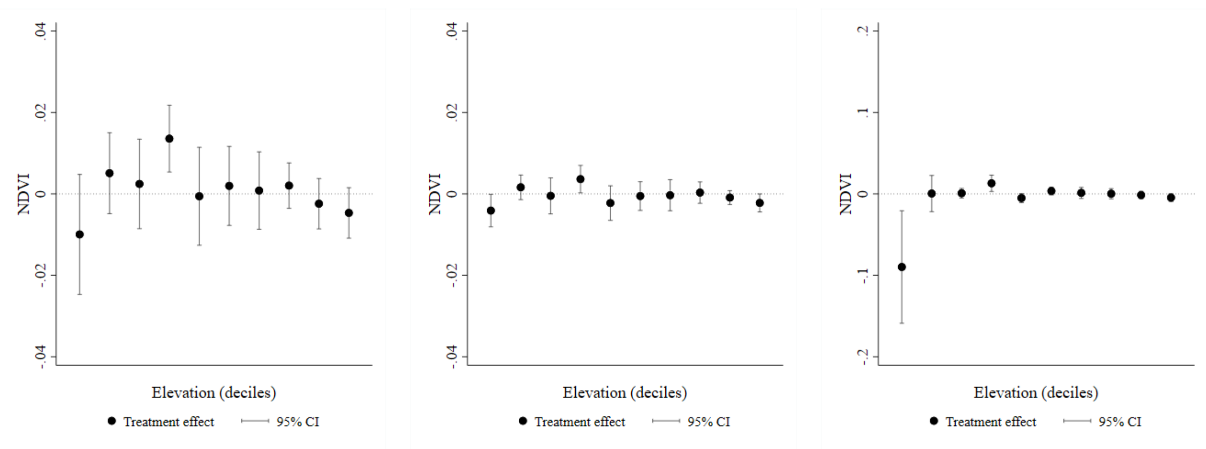
Table A4: Main results with standardized outcome

	Highlands			Lowlands		
	NDVI	NDVI	NDVI	NDVI	NDVI	NDVI
	(1)	(2)	(3)	(4)	(5)	(6)
Post × Treatment (dummy)	0.018			-0.019		
	(0.024)			(0.025)		
Post × Treatment (intensity)		-0.001			-0.010	
		(0.009)			(0.008)	
Post × Treatment (density)			0.011			0.000
			(0.017)			(0.019)
Woredas FE	✓	✓	✓	✓	✓	✓
Time FE	✓	✓	✓	✓	✓	✓
Time-varying controls	✓	✓	✓	✓	✓	✓
IP-weights	✓	✓	✓	✓	✓	✓
Mean dependent variable	0.00	0.00	0.00	-0.00	-0.00	-0.00
SD dependent variable	1.00	1.00	1.00	1.00	1.00	1.00
Observations	4606	4606	4606	2268	2268	2268
R-squared	0.91	0.91	0.91	0.96	0.96	0.96

Notes: Standard errors in parentheses are clustered at the level of the treatment (*woredas*). *** p<0.01, ** p<0.05, * p<0.1. See notes to Table 4 for other details.

Figure A3: Treatment effects by elevation deciles

(a) Treatment (dummy) (b) Treatment (intensity) (c) Treatment (density)



Source: Authors' elaboration.

Table A5: Main results with extended controls

	Highlands			Lowlands		
	NDVI	NDVI	NDVI	NDVI	NDVI	NDVI
	(1)	(2)	(3)	(4)	(5)	(6)
Post $\times$ Treatment (dummy)	0.001			-0.003		
	(0.002)			(0.003)		
Post $\times$ Treatment (intensity)		0.000			-0.001	
		(0.001)			(0.001)	
Post $\times$ Treatment (density)			0.001			-0.001
			(0.001)			(0.003)
Woredas FE	✓	✓	✓	✓	✓	✓
Time FE	✓	✓	✓	✓	✓	✓
Time-varying controls	✓	✓	✓	✓	✓	✓
IP-weights	✓	✓	✓	✓	✓	✓
Mean dependent variable	0.51	0.51	0.51	0.43	0.43	0.43
SD dependent variable	0.08	0.08	0.08	0.13	0.13	0.13
Observations	4606	4606	4606	2268	2268	2268
R-squared	0.91	0.91	0.91	0.96	0.96	0.96

Notes: Time varying controls include climatic variables (i.e. rainfall, temperature, and their respective quadratic terms), night time lights, and population density. Standard errors in parentheses are clustered at the level of the treatment (*woredas*). *** p<0.01, ** p<0.05, * p<0.1. See notes to Table 4 for other details.

Table A6: Oster bounds estimations

	Highlands			Lowlands		
	NDVI	NDVI	NDVI	NDVI	NDVI	NDVI
	(1)	(2)	(3)	(4)	(5)	(6)
Oster bounds $[\tilde{\beta}, \beta^{\star}(R_{max}, \delta)]$						
Post $\times$ Treatment (dummy)	[0.001,			[-0.002,		
	0.001]			-0.009]		
Post $\times$ Treatment (intensity)		[-0.000,			[-0.001,	
		0.003]			-0.003]	
Post $\times$ Treatment (density)			[0.001,			[-0.000,
			0.007]			-0.007]
List of controls in the full model						
Woredas FE	✓	✓	✓	✓	✓	✓
Time FE	✓	✓	✓	✓	✓	✓
Time-varying controls	✓	✓	✓	✓	✓	✓
IP-weights	✓	✓	✓	✓	✓	✓
Mean dependent variable	0.51	0.51	0.51	0.43	0.43	0.43
SD dependent variable	0.08	0.08	0.08	0.13	0.13	0.13
Observations	4606	4606	4606	2268	2268	2268
R-squared (full model)	0.91	0.91	0.91	0.96	0.96	0.96

Notes: $\tilde{\beta}$ corresponds to estimates using the full set of controls and $\beta^{\star}(R_{max}, \delta)$ to bias-adjusted estimates with $R_{max} = 1$ and $\delta = 1$ (see Oster (2019) for more details). See notes to Table 4 for other details.

Table A7: Main results with no restriction to the common support region

	Highlands			Lowlands		
	NDVI	NDVI	NDVI	NDVI	NDVI	NDVI
	(1)	(2)	(3)	(4)	(5)	(6)
Post × Treatment (dummy)	0.000			-0.002		
	(0.002)			(0.003)		
Post × Treatment (intensity)		0.000			-0.001	
		(0.001)			(0.001)	
Post × Treatment (density)			0.000			0.000
			(0.001)			(0.002)
Woredas FE	✓	✓	✓	✓	✓	✓
Time FE	✓	✓	✓	✓	✓	✓
Time-varying controls	✓	✓	✓	✓	✓	✓
IP-weights	✓	✓	✓	✓	✓	✓
Mean dependent variable	0.53	0.53	0.53	0.47	0.47	0.47
SD dependent variable	0.09	0.09	0.09	0.16	0.16	0.16
Observations	6594	6594	6594	2786	2786	2786
R-squared	0.93	0.93	0.93	0.97	0.97	0.97

Notes: Standard errors in parentheses are clustered at the level of the treatment (*woredas*). *** p<0.01, ** p<0.05, * p<0.1. See notes to Table 4 for other details.

Table A8: Main results keeping only the most treated woredas

	Highlands			Lowlands		
	NDVI	NDVI	NDVI	NDVI	NDVI	NDVI
	(1)	(2)	(3)	(4)	(5)	(6)
Post × Treatment (dummy)	0.003			-0.004		
	(0.002)			(0.004)		
Post × Treatment (intensity)		0.001			-0.002	
		(0.001)			(0.001)	
Post × Treatment (density)			0.001			0.002
			(0.001)			(0.002)
Woredas FE	✓	✓	✓	✓	✓	✓
Time FE	✓	✓	✓	✓	✓	✓
Time-varying controls	✓	✓	✓	✓	✓	✓
IP-weights	✓	✓	✓	✓	✓	✓
Mean dependent variable	0.51	0.51	0.51	0.43	0.43	0.43
SD dependent variable	0.08	0.08	0.08	0.13	0.13	0.13
Observations	3178	3178	3178	1372	1372	1372
R-squared	0.92	0.92	0.92	0.97	0.97	0.97

Notes: Sample of treated woredas restricted to the most treated woredas (i.e. woredas with above-average treatment density). Standard errors in parentheses are clustered at the level of the treatment (*woredas*). *** p<0.01, ** p<0.05, * p<0.1. See notes to Table 4 for other details.

Table A9: Impacts on land conservation

	Highlands			Lowlands		
	CA	CA	CA	CA	CA	CA
	(1)	(2)	(3)	(4)	(5)	(6)
Post × Treatment (dummy)	45.083			-17.478		
	(48.194)			(80.225)		
Post × Treatment (intensity)		22.056			-15.725	
		(22.454)			(29.789)	
Post × Treatment (density)			-55.170			96.178
			(33.846)			(68.089)
Woredas FE	✓	✓	✓	✓	✓	✓
Time FE	✓	✓	✓	✓	✓	✓
Time-varying controls	✓	✓	✓	✓	✓	✓
IP-weights	✓	✓	✓	✓	✓	✓
Mean dependent variable	0.51	0.51	0.51	0.43	0.43	0.43
SD dependent variable	0.08	0.08	0.08	0.13	0.13	0.13
Observations	658	658	658	324	324	324
R-squared	0.94	0.94	0.94	0.91	0.91	0.91

Notes: The outcome variable corresponds to the number of pixels cultivated at time t and still cultivated at time $t+1$ (with $t = 2001 \mid 2005$ and $t+1 = 2005 \mid 2013$). Standard errors in parentheses are clustered at the level of the treatment (*woredas*). *** p<0.01, ** p<0.05, * p<0.1. See notes to Table 4 for other details.

Table A10: Impacts on migration

	Highlands			Lowlands		
	Mig	Mig	Mig	Mig	Mig	Mig
	(1)	(2)	(3)	(4)	(5)	(6)
Post × Treatment (dummy)	0.015			-0.377		
	(0.013)			(0.311)		
Post × Treatment (intensity)		0.007			-0.114	
		(0.004)			(0.088)	
Post × Treatment (density)			0.010			-0.391
			(0.008)			(0.305)
Woredas FE	✓	✓	✓	✓	✓	✓
Time FE	✓	✓	✓	✓	✓	✓
Time-varying controls	✓	✓	✓	✓	✓	✓
IP-weights	✓	✓	✓	✓	✓	✓
Mean dependent variable	0.51	0.51	0.51	0.43	0.43	0.43
SD dependent variable	0.08	0.08	0.08	0.13	0.13	0.13
Observations	1568	1568	1568	664	664	664
R-squared	0.51	0.51	0.51	0.26	0.25	0.24

Notes: Authors' calculations based on IPUMS 2007 data. The outcome variable corresponds to the immigration rate (per 1,000 individuals) for woreda w at time t (with $t = \{2000, 2007\}$). Standard errors in parentheses are clustered at the level of the treatment (*woredas*). *** p<0.01, ** p<0.05, * p<0.1. See notes to Table 4 for other details.